*Electronic Letters on Computer Vision and Image Analysis 14(3):39-41, 2015*

# On a Distributed Video Surveillance System to Track Persons in Camera Networks

Niki Martinel

*Department of Mathematics and Computer Science, University of Udine, Via Delle Scienze 206, Udine, Italy*
*Advisor: Christian Micheloni*
*PhD thesis defended 9th April 2014, University of Udine*

Received 27th February 2015; accepted 27th July 2015

---

## 1 Contribution of the Thesis

In automated video surveillance applications, presenting the useful information to the human operators is a challenging task. Current systems usually require a prohibitive amount of human resources and lead to a quick decrease of the attention of the human operators through time, thus preventing them to catch the relevant events that may be worth to further investigate.

In addition, when monitoring a wide area, it becomes hard to deploy a network of video sensors such that there are enough overlapping FoVs to cover every point of the environment. This leads to the development of video surveillance systems (VSS) that provide partial area coverage. As a result, "blind-gaps" between camera FoVs are introduced. One of the most interesting problems which such "blind-gaps" bring in is to re-identify the persons moving across disjoint FoVs.

The contribution of the thesis * is two-fold. First, an advanced VSS is designed to display the proper task-dependent information to operators that are monitoring a wide area. In particular, the system helps operators in the task of tracking persons across camera views. This raised the need for a system capable of re-identify the subjects moving through disjoint FoVs. This leads to the second contribution of the thesis, a distributed approach to address the challenges of the person re-identification problem.

## 2 Adaptive Human Interface for VSS

The main goal of the proposed VSS is the development of an effective information visualization technique that properly displays only the most relevant cameras and information contents to simplify the operators' tracking tasks [1, 2]. The key idea is to visualize only most probable streams, i.e. those that will be involved with the motion of the tracked persons. Towards this goal we propose a novel dynamic organization, activation and switching of the User Interface (UI) elements based on the output of video analytics algorithms.

The first objective is to distill the volumes of monitoring information into a human manageable quantity. This is achieved by introducing an hand-off task between different camera views so that a single person can be tracked across different FoVs. The proposed camera planning algorithm uses geographical clues and exploits the predicted trajectories to build an accurate camera activation plan.

Correspondence to: <niki.martinel@uniud.it>

Recommended for acceptance by Jorge Bernal
DOI http://dx.doi.org/10.5565/rev/elcvia.717
ELCVIA ISSN:1577-5097
Published by Computer Vision Center / Universitat Autònoma de Barcelona, Barcelona, Spain

* Available at http://tinyurl.com/martinelthesis

The next and final objective is to present the filtered visual information to the operators such that they can take appropriate decisions in a limited amount of time. This is achieved by first exploiting the activation plan and tracking data such that only the proper streams are selected. Then, the selected video streams are displayed through a novel UI that allows the operators to focus only on a single view without requiring them to switch between monitors as well as UI elements. A map representation that exploits the *detail plus overview* technique is also introduced to make the task of inspecting the whole are less tough.

# 3 Distributed Person Re-Identification

In the adaptive human interface for VSS the main goal is to support surveillance operators in the task of tracking persons moving within the monitored environment. However, as a target exits a camera FoV it should be re-identified when it enters a different one, so as the task of tracking targets across camera FoV can be tackled. To address such re-identification problem, three main approaches are investigated. A distributed framework is also introduced to deal with the large amount of data that has to be shared across the whole network and which often saturates the communication and processing resources.

In the first approach, the re-identification is attacked by means of a discriminative signature based method [3, 4, 5, 6]. Given an image, the person body parts are detected using camera specific learned models, then local and global features are adopted to create a discriminative person signature. The signatures of two persons are matched using a weighted combination between local and global feature distances. While this method is effective for images that have similar appearance, it is not capable of dealing with the transformation of features that occur between disjoint cameras.

To tackle this issue, in the second approach, the nature of such transformation is investigated [7]. First, histogram-like features are extracted from local dense patches of the image. Then, for each feature extracted from two given images acquired by disjoint cameras, their transformation is captured by exploiting the principles of Dynamic Time Warping. The set of all such possible transformations form a function space. In such a function space, there are feasible and infeasible transformations. Feasible transformations are the ones that occur between features extracted from image pairs of the same person, infeasible transformations are the ones between features extracted from image pairs of different persons. The re-identification is carried out by learning the decision boundary that separates the feasible and infeasible transformations in the function space. A person is re-identified by classifying the transformation between image pairs as feasible or infeasible. This method strongly outperforms the previously proposed one when large color and illumination variations are present, however, due to the very high dimensionality of the function space it is computationally expensive.

To reduce such computational costs a tractable solution that can be later extended to perform the re-identification over the whole network is introduced [8, 9, 10, 11]. The idea is that as the features get transformed across cameras so are the differences between them. Thus, the distances between image features are exploited as new image pairs representations. This results in a much smaller space where the decision boundary separating the set of positive pairs (the two images are from the same person) and the set of negative pairs (the two images are from different persons) can be learnt. While being simple yet effective, this method still considers the re-identification as a camera pair process.

To consider the re-identification as a network process, a distributed framework is proposed [12]. The distributed network re-identification is achieved by first introducing a camera matching cost measure. Then, a derivation of the Distance Vector (DV) routing algorithm is adopted to route signatures to a specific subset of cameras in a priority fashion. Using an update rule similar to the one proposed by the original DV routing algorithm the network is capable of adapting through time.

# References

[1] N. Martinel, C. Micheloni, and C. Piciarelli, "Pre-Emptive camera activation for Video Surveillance HCI," in *ICIAP*, 2011.

[2] N. Martinel, C. Micheloni, C. Piciarelli, and G. Foresti, "Camera Selection for Adaptive Human-Computer Interface," *IEEE TSMC: Systems*, vol. 44, pp. 653–664, May 2014.

[3] N. Martinel and G. Foresti, "Multi-signature based person re-identification," *Electronics Letters*, vol. 48, no. 13, p. 765, 2012.

[4] N. Martinel and C. Micheloni, "Re-identify people in wide area camera network," in *CVPR Workshops*, 2012.

[5] M. Vernier, N. Martinel, C. Micheloni, and G. Foresti, "Remote feature learning for mobile re-identification," in *ICDSC*, 2013.

[6] N. Martinel and C. Micheloni, "Sparse Matching of Random Patches for Person Re-Identification," in *ICDSC*, 2014.

[7] N. Martinel, A. Das, C. Micheloni, and A. Roy-Chowdhury, "Re-Identification in the Function Space of Feature Warps," *IEEE TPAMI*, pp. 1–14, 2015.

[8] N. Martinel and C. Micheloni, "Person re-identification by modelling principal component analysis coefficients of image dissimilarities," *Electronics Letters*, vol. 50, pp. 1000–1001, July 2014.

[9] J. Garcia, N. Martinel, G. Foresti, A. Gardel, and C. Micheloni, "Person Orientation and Feature Distances Boost Re-Identification," in *ICPR*, 2014.

[10] N. Martinel, C. Micheloni, and C. Piciarelli, "Learning pairwise feature dissimilarities for person re-identification," in *ICDSC*, 2013.

[11] N. Martinel and C. Micheloni, "Classification of Local Eigen-Dissimilarities for Person Re-Identification," *IEEE Signal Processing Letters*, vol. 22, no. 4, pp. 455–459, 2015.

[12] N. Martinel, C. Micheloni, and C. Piciarelli, "Distributed Signature Fusion for Person Re-Identification," in *ICDSC*, 2012.